PROTOCOL

TITLE: **The Mere-measurement effect in patient-reported outcomes: A randomized control trial with speech pathology patients**

STUDY INITIATOR: Institute for Outcomes Research, MUW

MAIN AUTHOR: Preston Long[1]*, Valentin Ritschl[1,2]*, Melanie Ferk-Dornstauder[1], Tanja Stamm[1,2]*, Philipp Aichinger[3*]

[1] Medical University of Vienna, Centre for Medical Data Science, Institute of Outcomes Research, Vienna, Austria

[2] Ludwig Boltzmann Institute for Arthritis and Rehabilitation, Vienna, Austria

[3] Medical University of Vienna, Department of Otorhinolaryngology, Division of Phoniatrics-Logopedics, Speech and Hearing Science Lab, Vienna, Austria

*Contributed equally

Preface

The use of patient-reported outcome (PROs) have become increasingly commonplace across many healthcare settings over the past two decades. The value of PROs is now acknowledged by healthcare providers and patients alike. However, to date, little is known about the best practices for formulating PRO measures (PROMS), but even more specifically, the effect had on the responding patients as a result of item word choice, emotional valence, or frequency of use. That is, 1) does the positive or negative wording of items affect the patient's perspective on the latent variable, 2) is there a degree of subliminal influence or measurement effects on their behaviour resulting from exposure to PROs, and finally, 3) is such an effect amplified with repeated exposure?

1. FEASIBILITY STUDY PROTOCOL

1.1 BACKGROUND

There is currently sufficient evidence to suggest that attributes such as the wording of questions, their presentation order, the context where they are asked, and the item's social sensitivity (e.g. I do not abuse my prescriptions), have an effect on participant responses (Schwarz, 1999; Näher & Krumpal, 2012). This family of variables are most oft referred to as method effects. Method effects largely regard the phenomenon in which the presentation of an item affects the participant's response independent of the content in question. The end goal within this area of study is to correct for any response bias resulting from the items used. This research primarily concerns psychometrics. It is important to note that while method effects relate to the research questions proposed, it is not the key feature of interest. The ensuing inquiry will not simply focus directly on changes to item responses, but instead, on changes to patient perceptions and behaviours as a result of their exposure to the measures themselves. This has been coined the mere-measurement effect (Morwitz, Johnson, & Schmittlein, 1993).

Research supports the claim that asking certain questions can influence behaviours on the same topic (Sandberg & Conner, 2009; Godin et al., 2010). For example, a randomized-control trial was conducted to assess the impact of asking questions pertaining to blood donation habits of the participants. The researchers found that the subjects whom were asked about their habits were significantly more likely to donate blood than their control counter parts at the following blood-drive (Godin, Sheeran, Conner, & Germain, 2008). Other researchers found that when clinicians did not ask patients with Attention Deficit Hyperactivity Disorder (ADHD) to report whether they were currently medicated on a disease-related severity questionnaire, patients were more likely to report un-medicated symptomologies (Lineweaver et al., 2021). In other words, patients on ADHD medications responded more symptomatically similar to their un-medicated selves when their medication status was not included on the questionnaire. This suggests that the medication status item itself, may have mitigated ADHD symptoms by bringing the patient's treatment into a higher state of awareness/directing attention.

Thus, it may be that asking patients about their intent to have an operation increases the likelihood of them opting for it, or, that asking them about their exercise increases their physical activity level. If so, this may have significant implications for the effect of routinely collecting patient-reported outcomes (PROs), as well as for utilizing measurements as interventions. However, before value can be attached to such an effect, first the presence and strength must be evaluated in the patient-specific context.

To date, no mere-measurement effect factors have been investigated in the context of patient-reported outcomes measures (PROMs) and particularly not in regard to how they may influence patient perceptions and behaviours. Therefore, exploration of this family of

effects must be reduced to a few key anticipated features, namely, emotional valence (positive/negative) and frequency. For instance, would wording a cleft-lip question negatively, such as “I do not like my face”, actually decrease the patient’s preference for their own face, and, does repeated exposure to this question increase the magnitude of the effect.

This randomized controlled trial (RCT) review will be the first to the authors’ knowledge to examine if measuring patient- reported outcomes directly effects their corresponding behaviour, and if this effect can be attributed in part to item wording and/or frequency of exposure.

PROMs & Speech Disorders

Three recent reviews regarding patient-reported outcomes in speech pathologies were identified (Cohen & Hula, 2020; Francis et al., 2017; Slavych, Zraick & Ruleman, 2021). Francis et al. list a number of PRO measures that were designed for administration with subjects with any kind of voice disorders. The Evaluating Voice Disability-Quality of Life Questionaire (EVD-QOL) aims at assessing the general range of functional problems in voice disorders. The 30-items Voice Handicap Index (VHI) and its 10-items short version VHI-10 aim at providing a psychometrically robust voice diasability/handicap inventory. The Voice-Related Quality of Life (V-RQOL) aims at measuring the impact of voice problems on the quality of life. The Voice Activity and Participation Profile (VAPP) aims at identifying voice activity limitation and participation restriction separately. The Voice Symptom Scale (VoiSS) is an inventory of voice symptoms for assessing baseline pathology and response to change. The Communicative Participation Item Bank (CPIB) aims at assessing communication participation in all kinds of communication disorders. The Voice Self-Efficacy Questionnaire (VSEQ) monitors self-efficacy in individuals with self-declared voice problems before and after interventions. The Vocal Fatigue Handicap Questionnaire (VFHQ) and the Vocal Fatigue Index (VFI) aim at reflecting vocal fatigue.

Additional PROs include the Aging Voice Index, the Evaluation of the Ability to Sing Easily, the Glottal Function Index, the Linear Analog Scale of Assessment of Voice Quality, the Speech Disability Questionnaire, the Trans Woman Voice Questionnaire, the Vocal Cord Dysfunction Questionnaire, and the Vocal Tract Discomfort Scale Slavych, Zraick & Ruleman, 2021.

We have opted to use the CPIB and the VHI because these are well established and the former covers all kinds of communication disorders, while the VHI is using particularly negative wording in its original form.

1.2 STUDY RATIONALE

The purpose of the mere-measurement study is to identify the most advantageous way to measure outcomes for patients. This will be accomplished by collecting measures on the same topic at different time intervals and with different wording of questions. The findings of this study should help researchers and clinicians collect measurements from patients in the least burdensome and most beneficial manner possible.

The primary objective of the mere-measurement study is to assess the effects of different collection methods on the responders wellbeing.

The secondary objectives of this study are to:

- What frequency of collection is best for patients?
- What style of question wording is best for patients?

1.3 BENEFIT/RISK ASSESSMENT

No additional risks are anticipated through participation in the study than those that would be expected from daily life. The negative questionnaire group will receive an un-adapted, standardly collected questionnaire for speech pathologies, as such, we have not identified any potential risks associated with the participation in this study. Upon completion of your participation participants will receive a 50€ payment. In addition, participation in this study will provide insights to the scientific and medical community which in turn will help ensure we measure pathologies in the most beneficial manner possible. Potential benefits include societal contribution to healthcare and science, personal enrichment of the participants' treatment assessment and understanding of their wellbeing.

1.4 STUDY OBJECTIVES

The aim of this RCT is to identify and report the impact of collecting the PROMs controlling for topic content, and simply adjusting the frequency of collection and the emotional framing of the items. Thus, the researchers hope to shed light on the effects of measuring patient reported outcomes, as it relates to patient perceptions and behaviours on the item topics of inquiry.

Research Questions:

- Does the assigned positivity or negativity of an item effect patient perception or behaviour on the topic inquired?

- Does frequency of exposure to an item have an effect on patient perception or behaviour of the topic inquired?

Hypotheses:

H_1 = The more a patient reads and responds to questions about their speech condition the more their speech condition will be affected.

H_2 = The directionality of impact on a speech condition will be congruent with the emotional valence of the items used.

1.5 STUDY DESIGN

Once a subject has agreed to participate in this study, they will be asked to complete online questionnaires as well as record their voice reading a collection of random words 2-4 times.

This study follows a 2x2 design. Participants will be randomly assigned to one of four groups. Randomization will be accomplished by using the MUW randomizer. Participants will be stratified by education and speech disorder type. The healthy case control group will undergo case matching to ensure comparability across groups. Each group will receive two brief speech pathology severity questionnaires but at different frequencies, and with slight adjustments to the wording. Depending on your group, a subject will complete the questionnaires either two or four times, separated each time by one week. The questionnaires, in all groups, will take approximately 15 minutes to complete. Lastly, at the very start and once at the very end of the study, every participant will also be asked to record a brief 45 second audio file in the online survey as well as complete two additional psychological questionnaires.

The measures in this study cover the following tools (complete measures are uploaded in the ECS):

- A patient-reported outcome on voice health - Voice Handicap Index (VHI) or Voice Handicap Index Positively adapted (VHI-PA)
- A patient-reported outcome on speech health - Communicative Participation Item Bank (CPIB) or Communicative Participation Item Bank Positively adapted (CPIB-PA)
- A self-esteem stability scale - Self-Esteem Stability Scale (SESS)
- A self-esteem scale - Rosenberg Self-Esteem Scale (RSS)
- Two question measure of disease impact (2DB)
- A single-question measure of disease activity (1DS)
- Audio recording of text read aloud (ARec)

Start of study

The study start date will be the date of the first patient consenting to participate. The planned start date is **February 2024**

End of study

The end of the study will be the date from which the analysis has been completed on all data. The planned end of study date is **April 2025**

Length of study

This study will last **15 months**, including a recruitment period of **2 months.** Data of each participant will be collected over a period of 12 months.

The experimental/data collection portion of the study duration is only five weeks. The majority of the timeframe will be spent in recruitment, analysis and interpretation of the results.

Study population

All patients with Muscle tension dysphonia, Inducible laryngeal obstruction, Amyotrophic lateral sclerosis (ALS), suffering from a stroke or other brain injury/damage/trauma, (aphasia, dysarthria), and Parkinson's disease will be targeted for recruitment. There are no geographic location requirements. The participation schedule can be found below. Patients will be informed of the schedule in advance. A subject must complete any respective sampling epoch within 48 hours of the prescribed date in order to qualify as completed. Subjects must also have strong English skills and be technologically savvy to participate. Only adults will be included in the study.

Study Design Flow

Group 1*: High frequency-negative: 1 week between each collection*

1: Demographics, VHI, CPIB, SESS, RSS, 1DS, 1DB, ARec*

2: VHI, CPIB, SE, SSE, 1DS, 2DB, ARec

3: VHI, CPIB, SE, SSE, 1DS, 2DB, ARec

4: VHI, CPIB, SE, SSE, 1DS, 2DB, ARec

Group 2: *High frequency-positive: 1 week between each collection*

1: Demographics, VHI-PA, CPIB-PA, SESS, RSS, 1DS, 2DB, ARec*

2: VHI-PA, CPIB-PA, SE, SSE, 1DS, 2DB, ARec

3: VHI-PA, CPIB-PA, SE, SSE, 1DS, 2DB, ARec

4: VHI-PA, CPIB-PA, SE, SSE, 1DS, 2DB, ARec

Group 3: *Low frequency-negative: 2 weeks between each collection*

1: Demographics, VHI, CPIB, SESS, RSS, 1DS, 2DB, ARec*

2: VHI, CPIB, SE, SSE, 1DS, 2DB, ARec

Group 4*: Low frequency-positive: 2 weeks between each collection*

1: Demographics, VHI-PA, CPIB-PA, SESS, RSS, 1DS, 2DB, 1ARec*

2: VHI-PA, CPIB-PA, SE, SSE, 1DS, 2DB, 1ARec

Control: *(once at start, once at low frequency completion/final outcome collection, once at high frequency completion)*

1: Demographics, ARec x 3

Healthy Control: *(once at start, once at low frequency completion/ final outcome collection, once at high frequency completion)*

1: Demographics, ARec x 3

**completed prior to assessments*

Procedural step by step:

1. Probando advertises across their social media channels and website

2. Potential subjects log-in to Probando or create an account

3. Potential subjects are screened for inclusion and exclusion criteria in their profile

4. If they meet the criteria, and state permission to contact, their contact information will be passed on to the study investigators

5. The study investigators collect the informed consent via the online survey tool, following the patients review and signature

6. Each experimental subject will be randomly assigned into one of the four intervention groups; random assignment will not be used for the control or healthy group. The study investigators will handle all survey disbursement and participant communication over email or telephone.

7. As soon as the random assignment is complete, each subject will begin their respective trial spanning either two or four weeks. The study length is standardized and newly recruited participants will always start on the closest Wednesday to their entry, i.e. entered on a standard rolling basis. If subject's withdrawal from the study, their data will only be used up until the time point of their last complete submission but not in a full group analysis.

8. Payments will be made to participants once the last questionnaire has completed the trial.

1.6 PATIENT TARGET POPULATION

All patients will be required to have a qualifying disorder known to effect speech as identified below. Patients who meet the eligibility criteria will be invited to participate in the study by the researchers. Eligibility will be assessed prior to enrolment recruitment screening. AKH patients are not targeted for this study. All patients, regardless of geographic location around the world, can be recruited. The study is entirely online.

Subjects must meet the following inclusion criteria for study entry:

- Strong English skills*
- Technology savvy – able to complete online questionnaire
- Suffering from one of the following:

-Muscle tension dysphonia

-Inducible laryngeal obstruction

-Amyotrophic lateral sclerosis (ALS)

-Patients after a stroke or other brain injury/damage/trauma, (aphasia, dysarthria)

-Parkinson's disease

**English was chosen as the study language as a wider range of validated assessments exist in this language. The AI used for speech analysis is also best suited for and trained (66%) on English. Lastly, English is the most prolific language (native + secondary speakers) in the world which will facilitate the speed of recruitment.*

Subjects who meet any of the following criteria will be excluded from study entry:

- Under the age of 18

3.6.1 RATIONALE FOR SUBJECT POPULATION

The initial disease and/or problem areas, aforementioned, were chosen due to their known connection with speech disorders (Slavych, Zraick, & Ruleman, 2021). Speech disorders may be more susceptible to the mere-measurement effect as they are more closely influenced by an individual's psyche or mood state than many other conditions.

1.7 NUMBER OF SUBJECTS PARTICIPATING IN THE STUDY

We intend to include a total of 200 patients from across the acceptable condition types. They will be randomly separated into six groups including control and healthy control, with approximately 33 subjects per group. A formal power analysis was performed using G*power with parameters sourced from relevant seminal research (Conner et al., 2011; Godin et al., 2011). The following parameters were used for the sample size analysis: two-tailed directionality; effect size of 0.30; a priori computation, anticipate power of 0.95, and an error rate of 0.0023. Means difference tests between independent groups within the t-test family were utilized.

1.8 PARTICIPATING SITES

No additional sites will participate in this study.

1.9 RECRUITMENT PROCEDURE

Recruitment into the mere-measurement study will be performed supported by Probando. Probando is a previously approved patient recruitment services which specializes in online advertising. All recruitment materials are uploaded in the ECS. A description of the Probando recruitment service and procedures can be found as an attachment, as well as the recruitment materials for all subjects including healthy controls and a screenshot of the currently inactive landing page. As a part of their service, Probando uses the study inclusion and exclusion criteria to screen potential interested subjects. If they qualify and permit contact, their contact information will be provided to the MUW researchers for direct outreach. Therefore, all study timings, measures, and data will be distributed and collected only by the study researchers. Patients will be able to withdraw from the study at any time without providing a reason. To limit potential bias, an initial self-report of an official diagnosis is required to participate. Additionally, the patients will also confirm their diagnosis with a checkbox, for the purpose of minimizing the risk of malingering patients. If an individual is highly motivated to participate and does not suffer from a qualifying speech disorder they can still apply to participate as a healthy control. Furthermore, this sample would be more representative of a telemedicine community.

1.10 VARIABLES

Variables will consist of disease-specific PROMs, generic PROMs, audio recording of speech, and demographic variables. In order to minimize the non-interventional mere-measurement effect, the outcome measures were chosen as they are counter balanced between positive and negatively worded items.

Primary outcome variable

The primary outcome variable of this study is as follows:

- An audio recording of a standardized text, with the order of the paragraphs randomized in between every exposure. This text was adapted from the standard "Rainbow Passage" read aloud assessment with the field of speech pathologies (Dietsch et al., 2023). Artificial intelligence - ChatGPT - was used to recreate novel text of the same length and difficulty as the "Rainbow Passage" which is called "In the heart of a lush valley" (the text is uploaded in the ECS). An adapted version was opted for to eliminate the possibility of previous exposure.

Secondary variables

Secondary variables collected from patients include:

A patient-reported outcome on speech health - Communicative Participation Item Bank (CPIB)

- A self-esteem stability scale - Self-Esteem Stability Scale (SESS)
- A self-esteem scale - Rosenberg Self-Esteem Scale (RSS)
- Two question measure of disease impact (2DB)
- A single-question measure of disease activity (1DS)

Other variables of interest are:

- Demographics
- Qualifying disease or condition type

1.11 DATA SOURCES

Participants' and study data will be recorded in the study database. Study data will be collected at defined time points (see Section 1.5) using a single online format:

- SoSci survey (www.soscisurvey.de) is a routinely used, safe, online questionnaire distribution and data retrieval platform. Within this platform, five links will be created, one for each group. These will connect subjects to the questionnaires and to the instructions for and capturing of the audio file.

An email address for every participant will be required for communication and dissemination of participation credits. All individuals with access to participant personal information are obligated to respect their privacy and only use and disclose their information as described in the informed consent document.

Demographic information and mere-measurement study results are secured against unauthorized access. Security measures reduce the risk of unauthorized individuals accessing your personal information, but such risks cannot be entirely eliminated. All audio recordings will be stored only for the duration of the study. Upon closure, they will be deleted.

The information that we collect from the mere-measurement study will be kept confidential. Your personal information will be stored safely at MUW and pseudonymized as discussed in the previous section.

1.12 DATA COLLECTION

Online, subjects will complete a battery of questionnaires totalling approximately 15-20 minutes of time following the informed consent. Informed consent will be collected online via the SoSci survey tool. The following parameters will be applied to the virtual informed consent:

- The patient information including the data protection section is displayed above the checkbox. The patients are informed about what happens to their data, their benefits, and risks in sufficient detail but simple detail in this section.

- The checkbox says that patients have read and understood the text and that they agree.

- The patients will confirm their diagnosis with a checkbox, for the purpose of minimizing the risk of malingering patients

Depending on the group assigned, a participant will have to complete this battery either 2 or 4 times. The control and healthy group will complete only demographics questions and the audio recordings, no PROMs will be collected. The four intervention groups will be randomized using a digital randomization tool provided by the MUW (https://www.meduniwien.ac.at/randomizer/).

All patients, across all intervention groups, will complete the read-aloud activity prior to the first collection of PROMs. This will serve as the baseline for later comparison. Following the first recording, all subsequent recordings will be completed after completing the PROMs. An electronic bank transfer form (MUW standard bank form) will be emailed

to participants at the end of the study to complete the 50€ transfer. These electronic records will be immediately destroyed upon successful completion.

1.13 SUBJECT, STUDY, AND SITE DISCONTINUATION

3.13.1 SUBJECT WITHDRAWAL

Subjects have the right to withdraw their consent to data storage and collection at any time and without providing a reason. Expected reasons for withdrawal from the study may include, but are not limited to:

- Consent withdrawal
- Loss to follow-up
- Death

Where possible, the primary reason for withdrawal from the study should be documented. Subjects will not be followed up for any reason after consent has been withdrawn. Subjects who withdraw from the study will not be replaced. Partially completed data may be still included in the analysis if appropriate. Only participants who complete the study will receive payment. This will aid in subject motivation. There will also be additional bank transfer costs (16€ up to 37.40€) for individuals from countries outside the EU, thus, it is poor use of resources to make smaller payments

1.14 DATA MANAGEMENT

Patient-reported data will be captured using SoSci survey with an account owned by the Center for Outcomes research. These data will not be transferred outside of the MUW. Data transfers between the investigators of this study will be managed using procedures described in a jointly developed Data Transfer Plan. All data transferred between the two MUW centres, speech pathology and outcome research, will be identified only by a subject ID code, unless otherwise consented by patients and agreed upon in the Data Transfer Plan.

3.14.1 DATA QUALITY ASSURANCE

Patients will consent that pseudonymised patient data will be hosted and analysed at the MUW. MUW will not share the data with any third parties or collaborators. MUW regularly processes medical data and is subject to the GDPR requirements. MUW is responsible for data quality and will perform oversight of the data management of this study. Investigators and study coordinators will receive an initial training session on the protocol, study flow, study database, the survey tool, documentation, and expectations, and any applicable study processes. Access to data for secondary use cases will not be enabled.

3.14.2 SOURCE DATA DOCUMENTATION

Source documents (electronic) are those in which participant data are recorded and documented for the first time. The source documentation for this study will be:

- Participant demographic information and eligibility assessment
- Completed Informed Consent Forms
- Audio recordings
- Completed PROMs

1.15 STATISTICAL ANALYSES

Descriptive statistics and respective tables will be produced for every group which will include means, mediums, standard deviations, and size.

Primary outcome variable

The primary outcome variable of this study is as follows:

An audio recording of a standardized text, with the order of the paragraphs randomized in between every exposure. This text was adapted from the standard "Rainbow Passage" read aloud assessment with the field of speech pathologies (Dietsch et al., 2023). Artificial intelligence - ChatGPT - was used to recreate novel text of the same length and difficulty as the "Rainbow Passage" which is called "In the heart of a lush valley" (the text is uploaded in the ECS). An adapted version was opted for to eliminate the possibility of previous exposure.

The primary assessment parameter for the audio recordings is the number of words misspoken, as seen in the transcription. This parameter will undergo a secondary human assessment to ensure that the error was caused by the subject and not the AI. The length of the recordings will be used as a secondary factor. For the statistical analysis, we will prioritize accuracy over speed and perform a hierarchical hypothesis testing. The data and its residuals from the ANCOVA will be explored for normal distribution. If the data or residuals are non-normally distributed, we will use a nonparametric test or a longitudinal linear mixed model taking into account variance resulting from change over time.

Secondary variables

Secondary variables collected from patients include:

A patient-reported outcome on speech health - Communicative Participation Item Bank (CPIB)

- A self-esteem stability scale - Self-Esteem Stability Scale (SESS)
- A self-esteem scale - Rosenberg Self-Esteem Scale (RSS)
- A single-question measure of disease impact
- A single-question measure of disease activity

Other variables of interest are:

-Demographics

-Qualifying disease or condition type

Planned Analyses

Between Subjects Analysis:

All experimental groups in the 2x2 design will have their outcomes assessed in comparison to the control group as well as the healthy group at each collection point. This totals 24 tests which will require a family-wise error adjustment using the Bonferroni technique. The mean differences will be compared between each group respectively using an analysis of covariance (ANCOVA).

Within-Subjects Analysis:

All participants in every group will be compared to themselves over time. Each collection point, ranging from 2-4 depending on the group, will serve as an epoch. This can be completed through a single analysis using a repeated-measures analysis of covariance (RM-ANCOVA).

The following variables will be used as statistical controls in the analysis to ensure that any significant effect is appropriately assigned to the mere-measurement effect: *gender, age, response location, educational achievement, native language, and speech pathology disorder type.*

Post Hoc Analysis:

An exploratory post hoc analysis will be conducted comparing the scores of the interventions measures between positive and negative groups. This will provide insight on possible method effects resulting from the adaptation of the questionnaires. The mean

differences will be compared between both groups using an analysis of variance (ANOVA).

The primary outcome will be assessed using Automatic Speech Recognition (ASR) artificial intelligence software called Whisper (Radford et al., 2023). All analyses will be conducted on MUW computers and through MUW servers. No upload is necessary and the analyse will be ran in campus. All secondary outcomes measures will be scored and reported following their published procedures.

3.15.1 INTERIM AND FINAL ANALYSES AND TIMING OF ANALYSES

The following analyses for the study are planned:

- Assessment of the baseline audio recordings – week 1
- Analysis of low frequency completion – week 2
- Analysis of low frequency compared to control – week 3
- Analysis of high frequency completion – week 4
- Analysis of high frequency compared to control – week 5
- Analysis of high frequency compared to low – week 6
- Analysis of positive compared to negative items – month 2-5
- Analysis of positive compared and negative items compared to control - month 5-7
- Final analysis including all conditions and levels controlling for demographics – month 7-15

No mid-study decisions will be made based on the interim assessments. All comparisons will occur following final closure of the study data collection, i.e. following the last groups completion.

1.16 STUDY DOCUMENTATION AND MONITORING

The researchers will maintain adequate and accurate records to enable the conduct of the study to be fully documented, including but not limited to the protocol, protocol amendments, Informed Consent Forms, and documentation of IRB/EC. The study initiator shall ensure that the dataset and statistical programs used for generating the data included in the final study report are kept in electronic format and are available for auditing and inspection.

3.16.1 RETENTION OF RECORDS

Archiving at the study site will last for a minimum of five years after final study report or first publication of study results, whichever comes later; or according to local regulation. Records and documents pertaining to the conduct of this study must be retained by the

study initiator for at least 25 years after completion of the study, or for the length of time required by relevant national or local health authorities, whichever is longer. After that period of time, the documents may be destroyed, subject to local regulations. Bank details for participant pay will be deleted immediately following confirmation of payment receipt.

No records may be disposed of without the written approval of the study initiator. Written notification should be provided to the study initiator prior to transferring any records to another party or moving them to another location. All supporting functional parties will comply with the study initiator procedures regarding archiving and record management.

1.17 STUDY LIMITATIONS

The patient selection and the diagnostic or monitoring procedures are those applied per the usual treatment paradigm of the treating physician and not dictated by the protocol. As with all studies that require patients to self-report outcomes and behaviour, completeness and accuracy of reporting can be a concern. The data collection methods have been designed to be appropriate and accessible to the study population. Nevertheless, some errors in recording information and some missing information can be anticipated, particularly given the frequency of questionnaires required for data collection. As this study will collect PRO data through the use of devices, ability to use these devices is required and this may limit representativeness of the study sample. In summary, the design of this protocol presents inherent limitations that cannot be prevented and that are typical of non-interventional real-world studies. These studies have the potential for missing, inaccurate, or incomplete data. These limitations can result in methodological challenges in attributing causality to outcomes.

1.18 PROTECTION OF HUMAN SUBJECTS

3.18.1 COMPLIANCE WITH LAWS AND REGULATIONS

This study will be conducted in full conformance with the Declaration of Helsinki and the laws and regulations of the country in which the research is conducted. Data will be stored and handled according to the General Data Protection Regulation (GDPR). Patients will be made aware of their rights to consult or change their own data and of their possibilities to report any faults to the corresponding authorities.

3.18.2 INFORMED CONSENT

The Informed Consent Form will be provided via the online survey tool. The Consent Forms must be read, reviewed, and indicated online before start of documentation entering any data in the study database. By checking the consent form box, subjects confirm that they have been informed about the study procedures and agree to data collection. A digital copy of the Consent Form will be provided to the subject via a pop-up window with a download feature. The document titled ‘Information_questionnaire_first_page’ will be the

text located above the check-box informed consent; this is the first page the participants will see.

1.19 INSTITUTIONAL REVIEW BOARD OR ETHICS COMMITTEE

The mere-measurement study protocol, the Informed Consent Forms, and any study-related material will be submitted to the IRB/EC by the study initiators listed in this protocol and reviewed and approved by the IRB/EC before the study is initiated. In addition, any patient recruitment materials will be approved by the IRB/EC.

1.20 CONFIDENTIALITY

The study initiator maintains confidentiality standards by coding each subject enrolled in the study through assignment of a unique subject identification number. This means that identifiable data are not included in datasets that are transmitted to any study initiator's location. Subject medical information obtained by this study is confidential and will not be disclosed to third parties. The study initiator, including all listed investigators may use study data labelled with the patient ID numbers.

References:

Altmann, T., & Roth, M. (2018). The self-esteem stability scale (SESS) for cross-sectional direct assessment of self-esteem stability. Frontiers in psychology, 9, 91.

Baylor, C., Yorkston, K., Eadie, T., Kim, J., Chung, H., & Amtmann, D. (2013). The Communicative Participation Item Bank (CPIB): Item bank calibration and development of a disorder-generic short form.

Cohen, M. L., & Hula, W. D. (2020). Patient-reported outcomes and evidence-based practice in speech-language pathology. *American Journal of Speech-Language Pathology*, *29*(1), 357-370.

Conner, M., Godin, G., Norman, P., & Sheeran, P. (2011). Using the question-behavior effect to promote disease prevention behaviors: two randomized controlled trials. Health Psychology, 30(3), 300.

Dietsch, A. M., Mocarski, R., Hope, D. A., Woodruff, N., & McKelvey, M. (2023). Revisiting the Rainbow: Culturally Responsive Updates to a Standard Clinical Resource. American Journal of Speech-Language Pathology, 32(1), 377-380.

Fairbanks, G. (1960). Voice and articulation drillbook, 2nd edn. New York:

Harper & Row. pp124-139.

Francis, D. O., Daniero, J. J., Hovis, K. L., Sathe, N., Jacobson, B., Penson, D. F., ... & McPheeters, M. L. (2017). Voice-related patient-reported outcome measures: a systematic review of instrument development and validation. *Journal of Speech, Language, and Hearing Research*, *60*(1), 62-88.

Godin, G., Bélanger-Gravel, A., Amireault, S., Vohl, M. C., & Pérusse, L. (2011). The effect of mere-measurement of cognitions on physical activity behavior: a randomized controlled trial among overweight and obese individuals. International Journal of Behavioral Nutrition and Physical Activity, 8, 1-6.

Godin, G., Sheeran, P., Conner, M., Delage, G., Germain, M., Bélanger-Gravel, A., & Naccache, H. (2010). Which survey questions change behavior? Randomized controlled trial of mere measurement interventions. *Health Psychology, 29*(6), 636–644. https://doi.org/10.1037/a0021131

Godin, G., Sheeran, P., Conner, M., & Germain, M. (2008). Asking questions changes behavior: Mere measurement effects on frequency of blood donation. *Health Psychology, 27*(2), 179–184. https://doi.org/10.1037/0278-6133.27.2.179

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI) development and

validation. *American journal of speech-language pathology*, *6*(3), 66-70.

Lineweaver, T. T., Kercood, S., Gabor, A. J., Cervantes, J., Laine, A., & Baker, E. (2021). The effect of medication and question wording on self-reported symptoms and their accuracy in young adults with attention-deficit/hyperactivity disorder. British Journal of Clinical Psychology, 60(2), 252-269.

Morwitz, V. G., Johnson, E., & Schmittlein, D. (1993). Does measuring intent change behavior?. Journal of consumer research, 20(1), 46-61.

Näher, A.-F., & Krumpal, I. (2012). Asking sensitive questions: The impact of forgiving wording and question context on social desirability bias. *Quality & Quantity: International Journal of Methodology, 46*(5), 1601–1616. https://doi.org/10.1007/s11135-011-9469-2

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023, July). Robust speech recognition via large-scale weak supervision. In International Conference on Machine Learning (pp. 28492-28518). PMLR.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.

Sandberg, T., & Conner, M. (2009). A mere measurement effect for anticipated regret: Impacts on cervical screening attendance. *British Journal of Social Psychology*, *48*(2), 221-236.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*(2), 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Slavych, B. K., Zraick, R. I., & Ruleman, A. (2021). A systematic review of voice-related patient-reported outcome measures for use with adults. *Journal of Voice*.

van der Heijde D, Aletaha D, Carmona L, et al. 2014 Update of the EULAR standardised operating procedures for EULAR-endorsed recommendations. *Ann Rheum Dis* 2015;74(1):8-13. doi: 10.1136/annrheumdis-2014-206350